International Journal of Advanced Research in Engineering Technology & Science ISSN: 2394-2819

March-2016 Volume-3 Issue-3

Email: editor@ijarets.org www.ijarets.org

EXTRACTION OF HIDDEN AND META DATA FROM IMAGES

Karunanith. D
Department of Information Technology
Hindustan Institute of Technology & Science

Juvanna. I
Department of Information Technology
Hindustan Institute of Technology & Science

ABSTRACT: In the recent era, picture is an influential way to depict a meaning. It is a prerequisite to individualize reality from fantasy in a methodical way. Image is digitalizing the needed occurrence through the digital device. Pictorial representation of the scenarios plays a vital role in many real life happenings. In Cyber world, images have now become a crucial factor. The adversary side of Cyber world has cascade of consequences. Major crime in Cyber world happens through images like Morphing (image undergoes a gradual transformation into a slightly or completely new form through computer animation graphics) , Altering an image through Photoshop or other image editing software, Steganography (Embedding secret information in a text or image) , Leakage of digital documents and photograph. It is highly challenging for investigators, officials or judiciaries to identify these digital forgery and crime through image. The field of Image forensic is stimulated since deceptive photographs produced are eminent and commercial than the prototype. In this paper we are emphasizing Image Forensics to inspect the originality of the image and to identify the various frauds, damages, and threats cost to an individual or an organization or a country through various parameters of image metadata.

KEYWORDS: Meta Data, Error level Analysis, Digital Photos, and Image Forensics

INTRODUCTION

The standard of photography today in this modern era- the ingenious image alteration using Photoshop, the progression of digital camera, the persuasive role an image plays which makes it vital in our daily routine makes us apparent to neglect that the very first photograph was captured just 180 years ago. It is an irrefutable fact that the photography was a certain invention. In 1000AD, Pinhole camera also called as camera obscure was discovered by Alhazen (Ibn Al-Haythem), scholar in Optics. He was able to explain why the images were seen inverted. An optical device which is the axiom of photography and the digital camera consisting of a room or a box with a hole on one side is called as Camera Obscura. In 1826, Joseph Nicephore Niepce a French inventor fabricated the first photographic picture using camera obscura.[3] The heliographs of Neipce's are known as the primitive prototype for the modern images. Heliography is a photography technique using the Bitumen of Judea, which is then exposed to light. It lets the light to draw the picture. The picture titled View from the Window at Le Gras was taken at Neipce's estate which had set the stage for the evolution of photography in today's world. Images have become a crucial part in everyone's life. However, Exploitation of a picture has become distinct and apparent through the usage of nouvelle software. Eventually, endorsement of new intrigued technology helps in detecting the reformed image

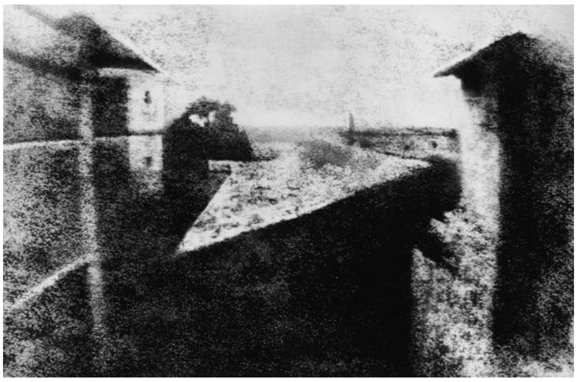


Figure 1 View from the Window at Le Gras INTRODUCTION

ROLE OF IMAGES IN CYBERWORLD

Cyberspace is an ideal medium in which communication over computer network works. During the years 1990s, when the use of internet, networking and digital communication was evolving the term ‘Cyberspace’ was able to portray many raw concepts and prodigy that were arising. Cyber-World can generally be classified as Countable and Uncountable. Uncountable computing is the world of inter-computer communication and Countable computing is a real or virtual world of data or information in Cyberspace. Images have now become an essential need in Cyber World. These being vital, major crimes in Cyber World develop due to Images. The issues faced are Morphing, Steganography, Leaked Photos, War Photos and Location of images.

A. **Morphing:-** When an image undergoes a gradual transformation into completely new form through computer animations or graphics then it is confined as Morphing. Commonly it is used to illustrate one person reformed to another .Typically such a description can be acquired through cross fading techniques.[2]

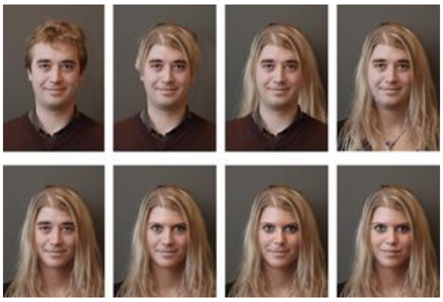


Figure 2 Methodology of Morphing

B. **War Photos:-** War Photos uses Morphing. An example can be taken to learn about it. A famous HBO show ‘Broadwalk Empire’ starring Jack Huston, who cast with half a face , a husky voice and barbarous speaking features Richard Harrow ,World War -I sniper [4] . He was horrifying when he go face sniped by an enemy sharpshooter. Here it seems like Hollywood had made him more alarming. [2]

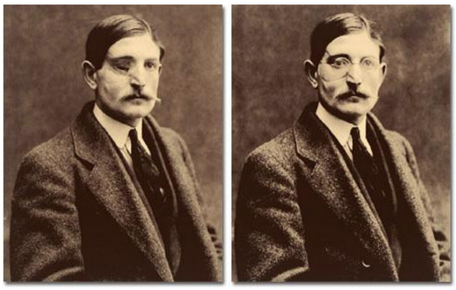


Figure 3 War Photo

C. **Steganography:-** Though steganography was formulated by the end of 15th century they were dated back to ancient Greece. In ancient Greece, people used wax tablet, message tattooed on shaved head and hair grew over it and shaved later to revel (Herodotus of Ancient Greece), invisible ink and much more. Steganography is embedding secret information on an image. Carrier is a transport medium which has a hidden message for the third party. [6] Below is the classification of Steganography Techniques (Adapted from Bauer 2002).

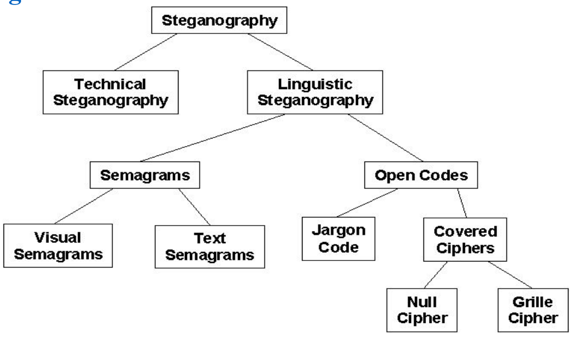


Figure 3 Classification of Steganography

D. Location of image:- In the era of Cyber World, images play a major role. There are fake snaps with morphed background going viral over the social media. In order to detect the location we can use geo tagging which identifies the longitude and latitude of that particular picture.

Figure 4 Image to depict Steganography

ANALYSIS OF IMAGE

Analyzing an Image addresses various disputes about an image being manipulated. [8]

1. Is the Image real or digitally refurbished?
2. If suppose the image is real, where and when was this taken?
3. If it was digitally reformed, what manipulations were done and how?

There are various approaches to analyse an image, which are as follows.

1. ***Observation:-*** No analysis tools are required. Direct observation can be used many a times to find hoaxed or fraudulent images. This is the easiest method using human analysis to extract data from an image.

Distinctive features that can be identified from an image are:

Reflection, Sharp highlights, Shadows, Items, Roots, Duplication, Scales,

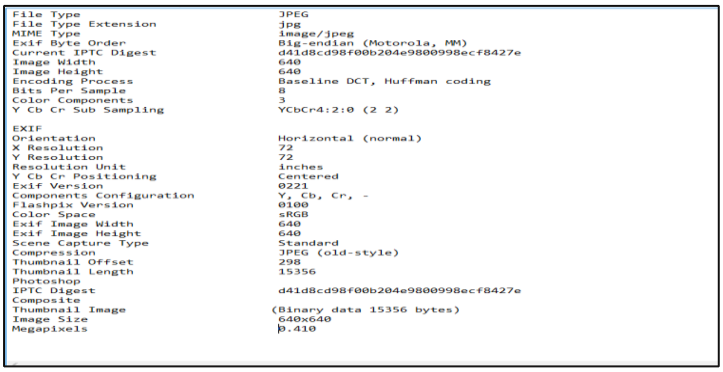
2. ***Basic Image refinement:-*** Familiar algorithms such as blurring, re-coloring, sharpening etc can be made more distinctive. Many photo editing tools use image refinement operation. It includes:

Sharpen, Blur, Color adjustment, Brightness and contrast

3. ***Image format analysis:-*** Formatting the image from one to other can be detected using format analysis tools.

IMAGE FORMAT ANALYSIS

Storing of images can be done in variety of formats such as RAW (Contains pixel data), JPEG, TIFF, PNG, GIF (Contains much information about the image). Any changes made in the format can be an efficient way to identify the changes. However, this paper concentrates only on JPEG images which have a well-bounded feature set that is modified when an image is altered. Meta data and JPEG Compression are the two main features which are discussed in this paper.



```
File Type                          JPEG
File Type Extension                jpg
MIME Type                          image/jpeg
Exif Byte Order                    Big-endian (Motorola, MM)
Current IPTC Digest                d41d8cd98f00b204e9800998ecf8427e
Image Width                        640
Image Height                       640
Encoding Process                   Baseline DCT, Huffman coding
Bits Per Sample                    8
Color Components                   3
Y Cb Cr Sub Sampling               YCbCr4:2:0 (2 2)

EXIF
Orientation                        Horizontal (normal)
X Resolution                       72
Y Resolution                       72
Resolution Unit                    inches
Y Cb Cr Positioning                Centered
Exif Version                       0221
Components Configuration           Y, Cb, Cr, -
Flashpix Version                   0100
Color Space                        sRGB
Exif Image Width                   640
Exif Image Height                  640
Scene Capture Type                 Standard
Compression                        JPEG (old-style)
Thumbnail Offset                   298
Thumbnail Length                   15356
Photoshop
IPTC Digest                        d41d8cd98f00b204e9800998ecf8427e
Composite
Thumbnail Image                   (Binary data 15356 bytes)
Image Size                         640x640
Megapixels                         0.410
```

Figure 5 Analysis of data

Error Level Analysis (ELA):- An image can be analysed by retrieving its metadata and re-saving the picture in lower quality. This is used to find extreme modifications. ELA evaluates the JPEG compression and creates a heat map which identifies error potential. It analyses the similar edges, textures and surfaces to produce a heat map. [**5**]

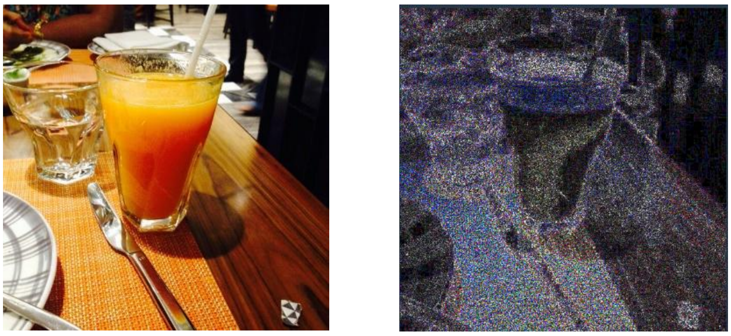


Figure 6 Heat map

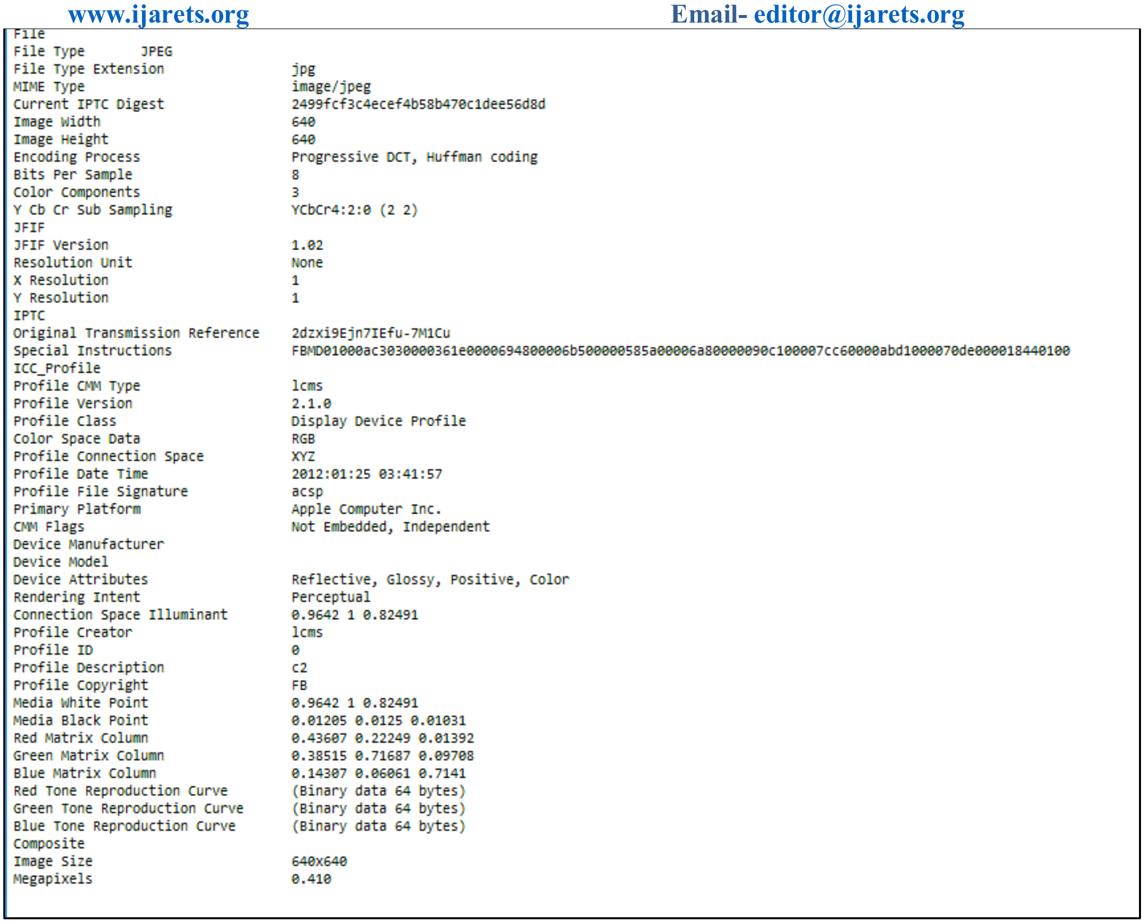
ANALYSIS OF METADATA

Data about data is called as Metadata. Images also contain metadata. These metadata contains all the information about the pictures which is mandatory for Image Forensics. Metadata is used to know how the file was created and handled. It provides information about the image like Make, Model, Software, Image Size, Time Stamps etc. The metadata is also be editable. But altering all metadata is not a simple process. Using Advanced Analysis alteration of metadata is also detectable. [7]

To analyze these metadata few Tools are used.

1- ***ExifTool:-*** An efficient cross platform tool which is used to derive metadata from almost any image format.
2- ***Exiv2:-*** An Open Source tool again used in retraining metadata of any format but not as much as ExifTool.
3- ***PNGmeta -*** An Open Source tool which analyses ONG files to give metadata.
4- ***Bio- Formats:-*** Metadata of microscopic images can be extracted.
5- ***Stegdetect:-*** An automatic tool which is used to find stenographic information in images.

Volume-3, Issue-3 March- 2016
www.ijarets.org Email- editor@ijarets.org

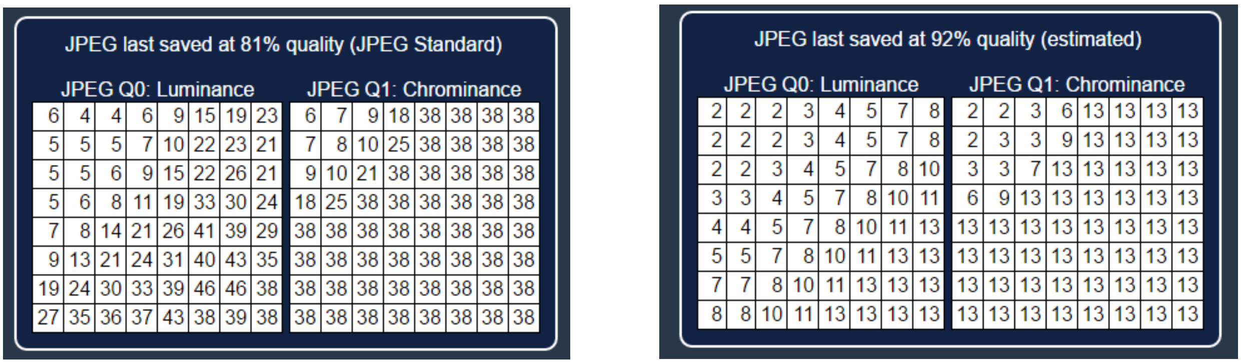


```
File
File Type       JPEG
File Type Extension                 jpg
MIME Type                           image/jpeg
Current IPTC Digest                 2499fcf3c4ecef4b58b470c1dee56d8d
Image Width                         640
Image Height                        640
Encoding Process                    Progressive DCT, Huffman coding
Bits Per Sample                     8
Color Components                    3
Y Cb Cr Sub Sampling                YCbCr4:2:0 (2 2)
JFIF
JFIF Version                        1.02
Resolution Unit                     None
X Resolution                        1
Y Resolution                        1
IPTC
Original Transmission Reference     2dzxi9Ejn7IEfu-7M1Cu
Special Instructions                FBMD01000ac3030000361e0000694800006b500000585a00006a80000090c100007cc60000abd1000070de000018440100
ICC_Profile
Profile CMM Type                    lcms
Profile Version                     2.1.0
Profile Class                       Display Device Profile
Color Space Data                    RGB
Profile Connection Space            XYZ
Profile Date Time                   2012:01:25 03:41:57
Profile File Signature              acsp
Primary Platform                    Apple Computer Inc.
CMM Flags                           Not Embedded, Independent
Device Manufacturer
Device Model
Device Attributes                   Reflective, Glossy, Positive, Color
Rendering Intent                    Perceptual
Connection Space Illuminant         0.9642 1 0.82491
Profile Creator                     lcms
Profile ID                          0
Profile Description                 c2
Profile Copyright                   FB
Media White Point                   0.9642 1 0.82491
Media Black Point                   0.01205 0.0125 0.01031
Red Matrix Column                   0.43607 0.22249 0.01392
Green Matrix Column                 0.38515 0.71687 0.09708
Blue Matrix Column                  0.14307 0.06061 0.7141
Red Tone Reproduction Curve         (Binary data 64 bytes)
Green Tone Reproduction Curve       (Binary data 64 bytes)
Blue Tone Reproduction Curve        (Binary data 64 bytes)
Composite
Image Size                          640x640
Megapixels                          0.410
```

Figure 7 Metadata of a downloaded image from facebook

Metadata of an Image varies with different parameters. In the above example, a random image saved in a computer was analysed which yielded the first metadata. The same image was then uploaded on Facebook and downloaded. The extracted metadata was very much different from the original picture.

JPEG Image Quality:- *Image* quality has to be selected while saving an image using an image editing tool. Higher the quality, bigger the image and better the resolution. Data loss can be calculated by knowing the percentage of image quality being saved. For example, an image which was saved under 98% will obviously result in minimal amount of data loss than an image being saved at 80% quality. More times the picture is saved or downloaded; the quality of the picture becomes less.



JPEG last saved at 81% quality (JPEG Standard)

JPEG Q0: Luminance								JPEG Q1: Chrominance							
6	4	4	6	9	15	19	23	6	7	9	18	38	38	38	38
5	5	5	7	10	22	23	21	7	8	10	25	38	38	38	38
5	5	6	9	15	22	26	21	9	10	21	38	38	38	38	38
5	6	8	11	19	33	30	24	18	25	38	38	38	38	38	38
7	8	14	21	26	41	39	29	38	38	38	38	38	38	38	38
9	13	21	24	31	40	43	35	38	38	38	38	38	38	38	38
19	24	30	33	39	46	46	38	38	38	38	38	38	38	38	38
27	35	36	37	43	38	39	38	38	38	38	38	38	38	38	38

JPEG last saved at 92% quality (estimated)

JPEG Q0: Luminance								JPEG Q1: Chrominance							
2	2	2	3	4	5	7	8	2	2	3	6	13	13	13	13
2	2	2	3	4	5	7	8	2	3	3	9	13	13	13	13
2	2	3	4	5	7	8	10	3	3	7	13	13	13	13	13
3	3	4	5	7	8	10	11	6	9	13	13	13	13	13	13
4	4	5	7	8	10	11	13	13	13	13	13	13	13	13	13
5	5	7	8	10	11	13	13	13	13	13	13	13	13	13	13
7	7	8	10	11	13	13	13	13	13	13	13	13	13	13	13
8	8	10	11	13	13	13	13	13	13	13	13	13	13	13	13

Figure 8 JPEG Compression of an original picture and after being downloaded from Facebook

CONCLUSION

Importance of analyzing an image grows stupendously as the pictures are widely manipulated which influences the opinion. Various analyzing strategies and tools are discussed in this paper. It is very unlikely that a modified image can escape from being identified, even if it fails one or two tests. Eventually, further analysis can be carried out to retrieve the fraudulence. The methods which are listed here are primitive and there are various advanced techniques which can further be implemented.

REFERENCES

1. http://blog.tumbhi.com ,by Tumbhi Team,18 Jan 2016.
2. http://www.cracked.com/article_20152_11-old-war-photographs-you-wont-believe-arent-photoshopped.html, By Eric Yosomono, January 3 2013
3. http://www.businessinsider.com/why-your-image-is-everything-12-2011?IR=T#dont-screw-up-like-this-either-10. By Judith Aquino and Kim Bhasin, Dec. 10, 2011, 12:35 PM.
4. http://twistedsifter.com/2012/02/famously-doctored-photographs/, [Source: fourandsix.com] feb 6, 2012.bbc.com
5. Gary C. Kessler, “An Overview of Steganography for the Computer Forensics Examiner”, February 2014,
6. Access Data. Forensic Toolkit product page [Online]. (December 29, 2003)
7. El-Khalil, R. Hydan [Online]. (December 30, 2003)